DOCUMENT RESUME

ED 058 883 LI 003 274

AUTHOR Helmuth, Nancy A.
TITLE The Use of Extracts in Information Services.
INSTITUTION ERIC Clearinghouse on Library and Information Sciences, Washington, D.C.
PUB DATE 71
NOTE 24p.; (30 References)

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Abstracting; *Abstracts; *Comparative Analysis; Guidelines; *Information Processing; Literature Reviews
IDENTIFIERS Extracting; *Extracts

ABSTRACT

It is the purpose of this paper to explore the use of extracts by reporting the results of a survey of the current acceptability and use of extracts, by a literature review and by suggesting guidelines for selecting extract material. The survey shows that extracts are currently being used in information services, especially for current awareness. Most consider extracting a time-saver. The methods for selecting material from a document are not well defined. A search of the literature produced only thirteen documents which deal with manually produced extracts. As reported in the literature, materials extracted include the author summaries, first and/or last paragraphs, selected sentences, the first page, and material selected by subject specialists from the entire text. The value of extracts as measured by how they lead the user to relevant documents was experimentally demonstrated. The purpose of the guidelines is to help the librarian and information specialist produce an informative extract. There was no substantial evidence that extracting is less costly than abstracting. The main criteria considers whether to spend the money to resay what the author has said or get a better understanding of the paper and use the author's words to describe it. (Author)

ED 058883

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

THE USE OF EXTRACTS IN INFORMATION SERVICES

By Nancy A. Helmuth
Information Scientist
ERIC Clearinghouse on Library and Information Sciences
1140 Connecticut Avenue, N.W.
Suite 804
Washington, D.C. 20036

1971

Introduction

To abstract or extract that is the question. Is it better to spend the time, money and intellectual effort necessary to put an author's ideas into an abstractor's words and thus create an original abstract, or will a careful selection of the author's own words best express the basic ideas of his paper?

It is the purpose of this paper to explore the use of extracts in information services by reporting the results of a survey of the current acceptability and use of extracts; by a literature review of the use, merits and methods of extracting; and by suggesting guidelines for selecting extract material from documents. Within the context of this paper, extracting is confined to manual procedures, and extracts are defined as: "Excerpts from a document used to represent the whole." (9). These excerpts consist of such material as: table of contents; all, or parts of, one or more paragraphs; the first page of the article; author summary; a selection of sentences from the full text of the document, or a combination of two or more of these items.

Survey

The attached questionnaire was sent to 390 librarians and information specialists considered to be the most likely to either be using extracts, or have an interest in the subject. From these, 76 useable responses were received. Thirty-three

LI 003 274

additional respondents declined to answer the questionnaire on the grounds that they did no extracting or abstracting.

The responses received were divided into three groups: the 25 respondents, who consider extracts suitable for all information services, were put into Group I; the 29, who consider extracts acceptable for some purposes, into Group II, and the 22, who do not believe extracts are ever acceptable, into Group III.

The significant results of the survey are presented below:

The purposes for which extracts are used, or considered suitable for use, by Groups I and II are given in Table 1. The prevalent specific uses of extracts are for current awareness (CA) and for selective dissemination of information (SDI).

Table 1: Uses of Extracts

Group	CA*	SDI+	Library Bulletin	Internal Reports	All Services	Not Specified
I	7	1	---	---	11	6
II	6	6	1	1	---	15
Total	13	7	1	1	11	21

*Current Awareness
+Selective Dissemination of Information

Only fourteen survey respondents do not consider extracting a time-saver, whereas twelve estimated that extracting saves 50% of the abstractor's (extractor's) time. The evaluation of extracting as a time-saver is found in Table 2.

Table 2: Estimated Time Saved by Using Extracts

Group	0	20	25	40	50	75	90	S*	Y***	NA+
I	4	1	--	1	5	1	1	3	7	2
II	3	--	1	1	7	2	--	3	6	6
III	7	1	--	--	--	--	--	--	4	10
Total	14	2	1	2	12	3	1	6	17	18

* S = saved time sometimes; no estimate given
** Y = saved time; no estimate given
\+ NA = no answer given

The most meaningful evaluation of any information service - though the hardest to obtain - comes from the user community. The use of extracts is no exception, as illustrated by Table 3 in which the majority of users gave no reaction to this method of information dissemination. However, user non-acceptance of extracts is reported in only six cases.

Table 3: User Reaction to Extracts

Group	Acceptable	No Reaction	Not Acceptable	No Answer
I	10	11	---	4
II	11	13	1	4
III	2	4	6	10
Total	23	28	7	18

When the length of the extract is set, the usual limit is 100 words, with the exception of one case in which a length of over 200 words was allowed, as shown in Table 4. When there is no specific word limit, the length of the extract is determined by the document content.

Table 4: Extract Length in Words

Group	*No Limit	50	50	100	200	200	No Answer
I	8	2	2	7	4	--	2
II	8	--	2	9	3	1	6
III	5	--	--	--	1	--	16
Total	21	2	4	16	8	1	24

* The limit is set by the document content with extracts kept as brief as possible was usually stipulated in this category.

The methods for selecting extract material from a document are not well defined. As seen in Table 5, only ten respondents have a written procedure for the process.

Table 5: A Written Procedure for Extracting is Used

Group	Yes	No	No Answer
I	4	18	3
II	2	22	5
III	4	7	11
Total	10	47	19

The document sections, or combination theory, which are considered most useful for extracting are listed in Table 6. Usually more than one section is used; a combination of the introduction and summary; a combination of the introduction, summary, and conclusions; or a combination of the summary and conclusions are most often named. The two sections most consistently used in combination with various other sections are the introduction and the summary.

Table 6: Document Section(s) Most Useful for Extract Material

Section(s)	Group I	Group II	Group III	Total
Introduction and Headings	1	--	--	1
Introduction and Conclusions	--	1	--	1
Introduction and Summary	5	3	--	8
Introduction, Preface, Summary, Key sentences and Results	1	1	--	2
Introduction, Summary and Conclusions	4	4	2	10
Introduction, Objectives and Summary	--	3	--	3
Introduction, Summary, Preface, Opening Chapter	--	1	--	1

Section(s)	Group I	II	III	Total
Introduction, Summary, Procedures and Definitions	1	--	1	2
Introduction, Summary, Methods, Results and Conclusions	--	1	2	3
Introduction, Foreword, Scope, Purpose, Summary and Table of Contents	1	--	--	1
Summary	--	2	1	3
Summary and Conclusions	3	2	--	5
First and Last Paragraphs	2	2	1	5
Methods and Results	1	--	--	1
Table of Contents	1	--	--	1
Discussion	--	1	--	1
Xerox of Top Page	--	1	--	1
No Special Preference	4	2	3	9
No Answer	1	5	12	18

Pinpointing the intellectual processes of extracting is very difficult. Nine of the respondents listed in Table 7 consider the intellectual processes for extracting to be the same as for abstracting and indexing. The majority, however, believe a subject specialist is required.

Table 7: Intellectual Processes Involved in Selecting Extract Material

Processes	Group I	II	III	Total
Inductive and Deductive Reasoning	1	--	--	1
Same as for Abstracting and Indexing	2	4	3	9
Scan for Key Ideas	7	6	3	16
Selective Judgement	3	7	1	11
*Subject Specialist Required	10	5	4	19
More Superficial then Abstracting	--	--	1	1
No Answer	2	7	13	22

* The prevalent quality required in the category of subject specialist is that the extractor know his user's needs and interests.

The strong feelings on both sides of the issue of extracts vs. abstracts are exemplified by some of the comments

elicited by this survey. For example, one respondent said that extracting did save the abstractor's time, but more of the user's time was required; another said it actually took longer to extract than to abstract, but extracts saved the users time.

Excerpts of some other comments elicited by the survey questionnaire from Group I include:

"Extracting, well-done, can avoid the pitfall of the abstract, which is the subtle insertion of the abstractor's bias. Also, extracts give the person using the information the word(s) from the horse's mouth."

"I have never felt that I need to abstract an article if the author has provided an acceptable substitute which I can extract. Inasmuch as my abstract bulletin is used by the readers to determine whether or not they wish to refer to the original articles, an indication of content is sufficient."

"Abstracts which say that...'The theory of such and such was developed and conclusions provided'...are of little value to anyone."..."Advise the reader of something of distinct value which he can use without going back and reading the article. This is the sort of technique that 'Extracting' should embrace."

"The small very specialized information service neither has the personnel nor time to spend on abstract creation. When commercial abstracting services are non-existent in your area, extraction may be the only answer."

Some of the comments from the Group II respondents are:

"We use the extract approach for the 'difficult' articles when it may not be obvious which wording should be used for a Current Awareness Bulletin abstract. Often the author's wording conveys the intended prupose and techniques of the research, and it would be a waste of time hunting for a new word description. Again, in subjects outside of the field of competence of the abstractor, it is more accurate to use the author's words to show the content of an article."

"I use extracts only when the report is written in such a manner that conciseness and clarity will not suffer. To use extracts in all cases would be more time-consuming than to start from scratch writing my own abstract."

"We are short on staff and long on material to be handled so we depend on the authors to abstract their own reports, and the library staff uses extracts in place of abstracts when we have to annotate."

"In my opinion, 'extracts' might be a useful and time-saving procedure for some purposes and with some materials. For example, if users are interested primarily in 'Results,' a sentence or two lifted from the results section of an article might be sufficient."

"Since our concern is eliciting information for our particular interests, we find extracting not sufficiently specific: in addition, we incorporate opinion, voiced by someone well acquainted with the subject matter. Therefore, our use of extracting instead of abstracting - plus - evaluation is only for expediency."

"The time saving" (by using extracts) "is mainly due to the fact that no dictation or longhand writing is required. Also, editing and proofreading can be dispensed with, or requires only clerical, not professional, effort."

"When extracting, one must be particularly careful to preserve the meaning of the information. A phrase may be taken out of context and combined with other phrases, resulting in something completely different in meaning, or at least in emphasis, from the original text."

"We have very rarely used an extract from a document to wholly substitute from an abstract. Invariably, an abstractor would edit such an extract to a greater or lesser extent in making it form part of the abstract."

"We have not been happy with extracting as a time saver."

Comments found in the Group II respondents include:

"Extracts, good ones, can only be made of extremely well-written articles or reports; those which have a well-defined Introduction, Conclusion, Summary. Most extracts are disjointed and sometimes miss vital associations."

"We find extracts inadequate to condense the entire document. However, systems which record specific data could very well use such a procedure."

"The average (author) abstract generally falls short in relating the pertinency of journal articles to the information needs of potential users. The author abstract is written from the viewpoint of only the author. A good abstractor writes an abstract from the viewpoint of the potential user."

"...the danger in extracting is that materials taken out of context can often be misleading and not present the true 'meat' of the article, and so, I say that extracting can be just as expensive as abstracting - for how can one extract from an article without scanning or reading the entire article."

"Any reduction in expense by the use of extracting instead of abstracting would probably be marginal. Both methods call for reading the documents...if the document has been read carefully and with understanding it is often quicker, easier, and produces better results, to write an abstract than to look for extracts which will adequately summarize the content."

"Extracting is lazy man's abstracting, and occasionally equivalent thereto."

The scarcity of literature pertaining to manual extracts or extracting is evidenced by the fact that only four respondents in Group I, three in Group II, and three in Group III, knew of any published information on this subject. The most cited were publications of the Battelle Memorial Institute, the best-known users of this method of information handling.

Literature Review

An exhausting, if not exhaustive, search of the literature produced only thirteen documents (1,5,6,10,11,12,13,15,17,18, 19, 23 and 25) which deal either directly or indirectly with manually produced extracts.

This is not indicative of the non-use of extracts however. S. Herner (12) found a widespread use of extracts by nine major indexing and abstracting publications incident to his investigation of abstract slanting. He reported that, in 207 cases in which the paper abstracted contained author abstracts or summaries, 46 were verbatim copies of the author abstracts or summaries, and 119 were variations of the author abstracts or summaries. In only 42 cases were there original abstracts,

that is, abstracts which bore no clear resemblance to the author abstracts or summaries. This was especially interesting in view of the response given to the survey letter of this study by one of the publications included in Mr. Herner's paper. The response stated that the publication did not consider extracts a satisfactory substitute for informative abstracts, that they do no extracting and that extracts would never be accepted by them.

The value of extracts as an information retrieval tool is reported in a study made by A. Kent et al. (13) who found that, to the motivated user, extracts of a combination of the first and last paragraphs of a document were at least as valuable as citations and abstracts for predicting the relevancy of a document.

In a study of abstract formation by a selection of sentences, G.J. Rath, et al.(18) compared the sentence selection by humans and machine. They found very little agreement between the human and machine methods in their selection of representative sentences. In a further study, A. Resnick (19) found that people selected the same sentences as representative of a document on a second trial in only 55% of the cases and that in only 64% of the cases could they correctly identify their previous selection. He felt this lack of reliability implied that there was no single set of representative sentences for an article. In a comparison of the usefulness of human, and machine-produced extracts with titles, and full text, G.J. Rath, et al. (17) found no major differences between the text, and machine and human-produced extracts in locating appropriate documents. However, titles alone led to a high incidence of "noise" i.e., documents which should have been rejected. The authors felt that the significantly high confidence in the human produced extracts over those by the

machine was due to their high literary quality and their sentence to sentence continuity.

J.H. Connor (5) suggests the use of xerox copies of the first page of articles as a first stage of a two-stage SDI system; the second stage being the full text. One of the respondents to this study reported using this method for SDI system. As pointed out by this user, the first page provides the title, author, source, and usually the author abstract if one is given. This seems to be providing a satisfactory low-cost SDI service.

No one in the field of information science has published on the development of the art of manual extracting to the degree attained at Battelle Memorial Institute. As early as 1956, the use of extracts in their Titanium Information Center was reported by R.W. Gibson and B.A. Lipetz (11). That Battelle is still using extracts is a good indication of the merits of this form of information handling. Other references to the extracting methods used at Battelle include: Guidelines for extracting for their Cooper Data Center (6); Applications of the Battelle Technique...by J.W. Murdock (15); and Qualitative Approach to Scientific Information Problems, by G.S. Simpson and J.W. Murdock (23). The most complete instructions for extractors is found in the as yet unpublished paper written for Battelle by C.A. Tippett (25). An indicative extract, by her definition, consists of just tables of contents, paragraphs or section headings, or initial paragraph or section sentences which only indicate the information contained in the document. Her suggestion for approaching a document to extract is the examination of the following parts in this order: Title page; author abstract - if given; summary and conclusions - sometimes contain more concise wording of the data presented; introduction - may be historical only; recommendations; charts and graphs; and finally the full text.

Another method of producing extracts briefly reported by J. Frome, et al. (10) consisted in having the authors and co-authors extract their own reports. The authors, response to this was reportedly favorable, but further information on this was not retrieved by my literature search.

Although there were no references to the cost of extracting found in the literature, two references concerned with abstracting costs are of interest. C.P. Bourne, et al. (3) states the gross unit cost of providing an abstracting service to be in the range of $5.00 to $30.00 per abstract. This is in agreement with the findings of R. Collison (4) who reported the cost of abstracting to be about $30.00 per article, including the bibliographic citation.

Guidelines

Ideally, extract material should be selected by a subject specialist as advocated by Battelle. However, since the ideal is rarely possible in a parctical situation, some guidelines might be helpful to the librarian or information specialist faced with the tasks of producing an informative extract.

The first matter of concern is what information should be extracted. Since the purpose of the extract is to substitute for an abstract, it seems reasonable for it to include it in the same kinds of information as/in an abstract. According to H. Borko, et al. (2), R. Collison (4), J.E. Rush, et al. (21), B.H. Weil (26) and Systems Development Corporation (24), an informative abstract should contain the: (1) purpose (scope), (2) methods, (3) results and (4) conclusions. And, just as important, H. Borko, et al. (2) Copper Data Center Bulletin (6) and B.H. Weil (26) recommended the exclusion of: (1) historical data; (2) repetition of previous results; (3) excessive verbiage /such as descriptions, explanations and speculations

and (4) repetition of the title.

Other more specific criteria for the extract content would be dictated by the user needs of the particular information system involved. The format and length of the extract are also matters for the individual information system to determine.

Having determined what information should be extracted, how long it can be and in what format, the next step is to examine the document. Careful scanning can usually determine whether the document belongs in the system, that is, whether it has information of value to the system's audience. When the document is included, the author abstract, if one is given, should be considered first. If it contains all the information needed, or, if it can easily be modified to fulfill all the requirements, it should be used, providing the author's abstract is informative, does not promise more than his paper delivers, and is oriented to the system's audience.

When there is no author abstract, or the author abstract is not suitable, the informative extract is taken from the text of the document. Fortunately, most research papers - the main items of information systems' input, are usually divided into the introduction, objectives, experimental part, discussion, conclusions, and summary. The introduction and discussion sections are of the least value for extracting since they usually contain the historical and background information, and the descriptions, explanations and speculations. However, if the extracter is not familiar with the subject matter of the document, these sections can give him valuable information and should be read. The objectives state the purpose of the work and the scope. The experimental section is examined for the method used to perform the work and the results - unless the latter are provided a section of their own. The conclusions contain the author's evaluation of his own work. The summary may give much of the information required for the extract in a relatively concise manner. In the absence of an author abstract, this is generally the most useful section for

the extractor. The captions beneath figures and tables should not be neglected for they can provide good extract material. The actual inclusion of appropriate graphs and figures is common practice at Battelle.

The burden of the actual selection of sentences to make up the extract is on the individual extractor. The article by J.E. Rush, et al. (21) is recommended for guidelines of specific things to watch for when selecting or rejecting sentences. The article is written for automatic abstracting but it applies equally well for the mannual extract. An extract can be "slanted" i.e., oriented to the system's users, by simply selecting the material most pertinent to this purpose. In many cases long sentences can be shortened by eliminating parenthetical phrases. Although the concern expressed about the dangers of changing the author's meaning ~~when~~ by extracting is probably overexaggerated ., the extractor must exercise extreme care not to alter the author's meaning or by what he selects or rejects . It seems more likely that the author's meaning would be changed if his words are also changed. Many, including G.J.C. Potter (16) believe it is best to use the original (author's) wording when abstracting. Further, H.P. Luhn (14) considers abstracts to be influenced by the abstractor's background, attitudes and opinions. Misinterpretation of the author's meaning due to such bias is less likely to occur in an extract. At the very least, the user can be confident that an extract does not contain more information than the document.

Unfortunately, not all of the documents that go into an information system are as easy to extract as the neatly arranged research paper. Speeches, letters to the editor, and certain other papers, may not contain a statement of purpose, a clear-cut method for arising at a conclusion or, for that matter, a real conclusion. Obviously, it would be impossible to produce an informative extract if the required information

(purpose, methods, results, and conclusions) is not in the document. In these cases, it is usually possible to extract a few sentences which indicate the document content, i.e., an indicative extract. The topic sentences - the first sentence or two of a paragraph - are usually the best for this purpose.

All documents are not easy to extract. Some will take longer than others and a few may not ever yield a good extract. These are old problems to an experienced abstracter and should not be insurmountable to the extractor. In these instances, the extract will not be better than the document it represents and the procurement of the document based on the extract will not dissapoint the user.

In the belief that actions speak louder than words, I conclude these guidelines with three examples of extracts I have done. These extracts represent documents from the biological sciences, the physical sciences, and library and information sciences.* The extracts are limited to approximately 200 words - an acceptable length for abstracts.

Example 1.

Experimentation With Computer-Assisted Indexing: American Documentation, Volume 20. Hines, Theodore C.; Harris, Jessica L. and Colverd, Martin, Journal of the American Society for Information Science 21(6):402-405, Nov-Dec 1970.

> The index to the 1969 volume of American Documentation was the final output to a medium-scale test of a system and programs for computer-assisted indexing developed at Columbia University School of Library Service. The term "computer-assisted" is used to distinguish the system from more automatic title (or text) derivative methods. The subject terms are assigned by a human indexer and all other access points used are program-retrieved from the input data. Only procedures for indexing journal articles are described but other materials would be handled similarly. The basic steps in the process include: (1) Marking and indexing journal issues; (2) Recording, in machine-readable

*Permission given by the American Association for the Advancement of Science to extract from their publication Science and from the American Society for Information Science to extract from the Journal of the American Society for Information Science is greatly appreciated.

form, journal information, article information, and subject or added entries assigned by the indexer; (3) A program job step that creates full entries and sort keys for each access point specified for the index in question; (4) A three-job program segment that arranges the sort keys and (5) A final program that takes entries and sort keys as input and produces a finished, paged printout of camera-ready copy of the index. The output provides author, title, and subject entries and the complete bibliographic citation at each access point. The system is demonstrated to be economic in comparison to currently used manual methods.

Example 2.

Carbon Monoxide: Its Role in Photochemical Smog Formation. Westberg, Karl and Cohen, Norman, Science 171 (3975): 1013-1015, Mar 12, 1971.

Experimental evidence is presented for the effect of CO and the oxidation of NO in polluted atmospheres. Experiments were performed in a 7.68-m^3, Teflon-lined, constant-temperature, stirred smog chamber. Approximately 3 parts per million (ppm) of isobutene, 1.5 ppm of NO or NO_2, and varying amounts of CO were mixed with an atmosphere of pure air and irradiated with light approximating the intensity and spectral characteristics of sunlight. High pollutant concentrations and the relatively high volume-to-surface ratio (V/S=0.227m) of the smog chamber minimized the importance of wall reactions. Continuous measurements were made of the concentrations of O_3, NO_2, $NO+NO_2$, CO and isobutene. The results show that CO accelerates the oxidation of NO to NO_2 and the rate of ozone formation. However, when the ozone concentration becomes greater than about 0.6 ppm, CO decelerates the further oxidation of NO to NO_2 and the rate of ozone formation. The following conclusions are established: First, CO is not inert in the production of photochemical smog. Second, OH is implicated as a major chain carrier in the smog-producing system. Finally, qualitatively it seems that the presence of CO will accelerate the earlymorning conversion of NO to NO_2, hastening the appearance of oxidants.

Example 3.

Perinatal Undernutrition: Accumulation of Catecholamines in Rat Brain. Shoemaker, William J. and Wurtman, Richard J., Science 171 (3975): 1017-1019, Mar 12, 1971.

> In order to study the effects of early undernutrition on a specific property of brain neurons we have measured substances found exclusively within those cells in the brains of rats experimentally malnourished from birth to weaning. Our data indicates that 24 days after birth, such brains contain 25 to 30 percent less of the neurotransmitter norepinephrine than those of control animals do. Amounts of dopamine in the brain are also depressed but the activity of the catecholamine-synthesizing enzyme tyrosine hydroxylase is significantly elevated. Four experimental groups of young were generated: (i) pups born to control mothers and nursed by control mothers (C-C); (ii) pups born to control mothers and nursed by deprived mothers (C-D); (iii) pups born to deprived mothers and nursed by control mothers (D-C); and (iv) pups born to and nursed by deprived mothers (D-D). Prenatal deprivation did not depress amounts of norepinephrine in the brain among animals suckling control mothers (D-C group); however, it did magnify the effect of postnatal malnutrition in the D-D group. Amounts of dopamine in the brain were also lower 24 days after birth in rats nursed by deprived mothers. Our results show that when rats are undernourished prior to weaning, the accumulation of brain norepinephrine and dopamine is impaired.

Summary

In essence, the survey shows that extracts are currently being used in information services, especially for current awareness. (Table 1). Most consider extracting a time-saver; twelve estimate a savings of 50% of their time is possible by extracting instead of abstracting (Table 2). The acceptance of extracts by the user community appears to be rather passive which is not unusual for any information service. (Table 3). The length of extracts is approximately the same as for most abstracts - 100 to 200 words (Table 4). The actual methods used for selecting material to be extracted is the most vague area.

Most extract users have no written procedures (Table 5) and there is a great difference of opinion as to which document sections or combination of sections are most useful for extracting (Table 6). The belief that a subject specialist is required for extracting is held by many of the survey respondents (Table 7). The very definite opinions of the librarians and infomation specialists either for or against the use of extracts in information services are brought out by the excerpts of some of the survey comments.

As reported in the literature, extracts have been used by Battelle Memorial Institute as well as by other indexing and abstracting organizations. The materials extracted include the author summaries, first and/or last paragraphs, selected sentences, the first page, and material selected by subject specialists from the entire text. For indicative extracts, such materials as tables of contents, section headings, and initial paragraph or section sentences are useful extract material.

The value of extracts is best measured by how they lead the user to relevant documents. This value was at least experimentally demonstrated by A. Kent et al. (13) and by G.J. Rath et al. (17).

The purpose of the guidelines is to help the librarian and information specialist produce an informative extract, to determine what material should go into the extract and in what sections of the document this information is most likely to be found.

Besides the instructional material, the three examples of extracts given illustrate what can be expected from this method of information processing. The extract of this paper can serve as a fourth example.

There was no substantial evidence that extracting is less costly than abstracting. The main criteria seems to depend on whether it is better to spend the money (time) to resay what the author has already said, or to get a better understanding of the paper and use the author's own words to describe it.

For an example of a working information system which uses extracts and structured thesaurus index terms, the paper by W.T. Black and W.S. Lyman (1) is recommended.

Acknowledgements

For the reviews by Mr. D.M. Liston and Mr. S.M. Newman, for the editorial help of Mr. S.M. Newman and Ms. Alice Billingsley and for the support and helpful suggestions of Mr. J.I. Smith, I am sincerely grateful.

ATTACHMENT

ERIC CLEARINGHOUSE ON LIBRARY AND INFORMATION SCIENCES

Please complete and return this questionnaire to: ERIC/CLIS, 1140 Connecticut Avenue, N.W., Suite 804, Washington, D.C. 20036. ATTN: Miss Nancy Helmuth

1. Do you consider extracts a satisfactory substitute for writing abstracts or annotations?

2. What intellectual processes are involved in selecting the extract?

3. Do you have a written procedure? If so, would you state it briefly?

4. Do you use extracts only for special purposes, such as SDI or Current Awareness services, or for all services requiring abstracts or annotations?

5. How do you limit the extract length? 50 words, 100 words, 200 words, other _____?

6. Are the extracts always taken from just one paragraph or do you sometimes select parts of paragraphs from different locations in the text?

7. Have you had any reaction from your users concerning the adequacy of extracts?

8. Is extracting really a time saver? Could you estimate the percentage of time saved, if any?

9. If you know of any publications covering extracting, please give us references to them.

10. What, if any, particular section or part of the article do you usually find appropriate extract material?

Additional Remarks:

BIBLIOGRAPHY

1. Black, William T. and Lyman, W. Stuart
   The Copper Data Center -- A Total-Access System. Paper presented at the Forum on the Management of Information Analysis Centers, Gaithersburg, Maryland May 17-19, 1971, 5pp.

2. Borko, Harold and Chatman, Seymour
   Criteria for Acceptable Abstracts: A Survey of Abstractors' Instructions. American Documentation 14(2):149-160, Apr 1963.

3. Bourne, Charles P., North, Neanne B., and Kasson, Madeleine S.
   Abstracting and Indexing Rates and Costs: A Literature Review. ERIC Clearinghouse for Library and Information Science, May 1970, 9pp. ED 043 789, MF-$0.65; HC-$3.29. EDRS, Leasco Information Products Co., P.O. Drawer O, Bethesda, Maryland 20014

4. Collison, Robert
   Abstracts and Abstracting Services. American Bibliographical Center - Clio Press Rivera Campus, 2040 Alameda Padre Serra, Santa Barbara, California, Oxford, England, 1971, 122pp.

5. Connor, Judith Hall
   Selective Dissemination of Information: A Review of the Literature and the Issues.
   In: Elias, Arthur W., ed. Key Papers in Information Science, The American Society for Information Science, 1971, pp. 135-153.

6. Copper Data Center Bulletin
   Guidelines for Industry Specialists in Extracting. Document on Copper Technology. Copper Data Center Bulletin, Nov 1968, 2pp.

7. Costello, John C., Jr.
   Training Manual and Workbook for Use in Abstracting and Coordinate Indexing. Battelle Memorial Institute, Columbus, Ohio, Revised 1964, 125pp.

8. DeLucia, Al
   Index-Abstract Evaluation and Design. American Documentation 15(2):121-5, 1964.

9. Evans, Renee C., Comp.
   Tutorial Glossary of Documentation Terms. Compiled for the 29th Annual Meeting of the American Documentation Institute, Oct 307, 1966, Santa Monica, California, 78pp.

10. Frome, Julius; Caponio, Joseph F., and Brown, Virgil
Author Extracts in Addition to Technical Reports: An Aid Toward Timely Dissemination of Technical Reports. In: Automation and Scientific Communication, Topic I: Generation of Information, 1963 ADI Annual Meeting. p73.

11. Gibson, Robert W., Jr. and Lipetz, Ben-Ami
New Look in Manual Methods. Special Libraries 47(3): 106-113, Mar 1956.

12. Herner, Saul
Subject Slanting in Scientific Abstracting Publications. Proceedings of the International Conference on Scientific Information 1960, pp.407-427.

13. Kent, A., Belzer, J.; Kurfeerst, M.; Dym, E.D.; Shirley, D.L., and Bose, A.
Relevance Predictability in Information Retrieval Systems. Methods of Information in Medicine 6(2): 45-51,Apr 1971.

14. Luhn, H.P.
The Automatic Creation of Literature Abstracts. In: Elias, Arthur W., ed. Key Papers in Information Science, The American Society for Information Science, 1971, pp87-93.

15. Murdock, J.W.
Application of Battelle Technique to the Operation of Information Centers.
Journal of Chemical Documentation 2:126-127, 1962.

16. Potter, G.J.C.
In Singer, T.E.R. Ed. Information and Abstracting Communication Practice in Industry. 1958. pp.281-292.

17. Rath, G.J., Resnick, A. and Savage, T.R.
Comparisons of Four Types of Lexical Indicators of Content. American Documentation 12(2):126-130, Apr 1971

18. Rath, G.J., Resnick, A. and Savage, T.R.
The Formation of Abstracts by the Selection of Sentences. Part I. Sentence Selection by Men and Machines. American Documentation 12(2):141-143 Apr 1961

19. Resnick, A.
The Formation of Abstracts by the Selection of Sentences. Part II. The Reliability of People in Selecting Sentences. American Documentation 12(2):141-143, Apr 1961

20. Resnick, A. and Savage, T.R.
The Consistency of Human Judgements of Relevancy.
American Documentation 15(2):93-95, Apr 1961.

21. Rush, J.E., Salvador, R. and Zamora, A.
Automatic Abstracting and Indexing, II Production of Indicative Abstracts by Application of Contextual Inference and Syntactic Coherence Criteria.
Journal of the American Society for Information Science, 22(4):260-274, Jul-Aug 1971

22. Sieburg, J.
Automatic Abstraction of Legal Information, Datamation, 12 pp.63-65, Nov 1966.

23. Simpson, G.S. Jr. and Murdock, J.W.
Qualitative Approach to Scientific Information Problems.
Battelle Technical Review, Nov 1960, 5pp.

24. Systems Development Corporation
A System Study of Abstracting and Indexing in the United States. Dec 16, 1966, PB 174 249, 228pp, (NTIS, Springfield, Va. 22052)

25. Tippett, Carol A.
Introduction to Extracting and Clue-Wording,
Battelle Memorial Institute, Jul 30, 1970, 59pp. (unpublished)

26. Weil, Ben H.
Standards for Writing Abstracts. Journal of the American Society for Information Sciences 21(5):351-357, Sept-Oct. 1970

27. Weil, Ben. H.; Zarember, I., and Owen, H.
Technical-Abstracting Fundamentals. I. Introduction.
Journal of Chemical Documentation 3(3):86-89, 1963

28. Weil, Ben H; Zarember, I., and Owen, H.
Technical-Abstracting Fundamentals II. Writing Principles and Practices. Journal of Chemical Documentation 3(3):125-132 , 1963.

29. Weil,Ben H.; Owen, H. and Zarember, I.
Technical-Abstracting Fundamentals III. Publishing Abstracts in Primary Journals. Journal of Chemical Documentation 3(3):132-136 , 1963.

30. Welt, Isaac D.
Abstracting for Technical Information Centers. Technical Information Center Administration 3:27-43, 1967

EXTRACT

THE USE OF EXTRACTS IN INFORMATION SERVICES

It is the purpose of this paper to explore the use of extracts by reporting the results of a survey of the current acceptability and use of extracts, by a literature review and by suggesting guidelines for selecting extract material. The survey shows that extracts are currently being used in information services, especially for current awareness. Most consider extracting a time-saver. The methods for selecting material from a document are not well defined. A search of the literature produced only thirteen documents which deal with manually produced extracts. As reported in the literature, materials extracted include the author summaries, first and/or last paragraphs, selected sentences, the first page, and material selected by subject specialists from the entire text. The value of extracts as measured by how they lead the user to relevant documents was experimentally demonstrated. The purpose of the guidelines is to help the librarian and information specialist produce an informative extract.. There was no substantial evidence that extracting is less costly than abstracting. The main criteria considers whether to spend the money to resay what the author has said or get a better understanding of the paper and use the author's words to describe it.